

# Interactive Visualization of Video Metadata

Mark Derthick

Human Computer Interaction Institute  
Carnegie Mellon University  
Pittsburgh, PA, USA 15213  
1 412 268-8812

mad@cs.cmu.edu

## ABSTRACT

Much current research on digital libraries focuses on named entity extraction and transformation into structured information. Examples include entities like events, people, and places, and attributes like birth date or latitude. This video demonstration illustrates the potential for finding relationships among entities extracted from 50,000 news segments from CMU's Informedia Digital Video Library. A visual query language is used to specify relationships among entities. Data populate the query structure, which becomes an interface for exploration that gives continuous feedback in the form of visualizations of summary statistics. The target user is a data analyst familiar with the domain from which the entities come, but not a computer scientist.

## Categories and Subject Descriptors

H.5.2 [User Interfaces]: Graphical user interfaces (GUI), Interaction styles.

## General Terms

Algorithms, Human Factors

## Keywords

Information Visualization.

## 1. METADATA EXTRACTION

CMU's Informedia News on Demand project has recorded and processed thousands of hours of news video from CNN and other sources [1]. A transcript of the audio is obtained from closed-caption information or speech recognition. Each word is looked up in a gazetteer of geographical locations. Any locations found are linked to the video segment, and the number of times the location is mentioned is recorded in a database of metadata. The latitude and longitude, country, and continent are also added.

Separately, the video frames are analyzed by a face recognition algorithm and optical character recognition. Frames containing a name and a face give rise to a "named face" entity. The person's title is extracted by OCR or looked up in a biographical dictionary. The latter source also adds nationality and dates of birth and death.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

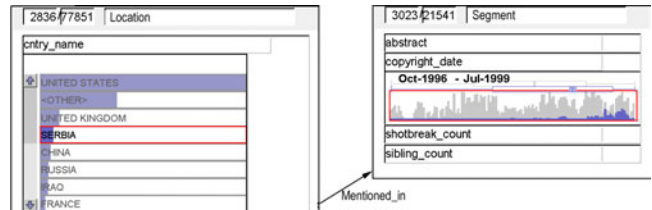
JCDL '01, June 24-28, 2001, Roanoke, VA, USA.

Copyright 2001 ACM 1-58113-000-0/00/0000...\$5.00.

The Informedia database subset shown contains about 50K video segments, 1500 named face occurrences of 200 distinct names, and 80K occurrences of 1800 distinct locations.

## 2. ENTITY VISUALIZATION

The visual query language allows database-style joins between entities of different types, Dynamic Query filtering of attribute values, visualization of conditional attribute value distributions with histograms, and drill down to individual entities [2]. For instance, the distribution of the countries of the geographical locations in the figure below shows the dominance of the US, CNN's home country, but also concentrations that reflect specific events, as in the case of Serbia. By clicking on histograms bars in the left box, the user can focus on locations in Serbia. This restriction propagates along the Mentioned\_in relationship to the right box. There are 21,541 segments that mention some location, of which 3023 mention a location in Serbia. The dark conditional distribution of the segment copyright\_date attribute shows that most of the Serbia segments were recorded in 1998-1999, while the light unconditional distribution shows that segments overall are distributed more uniformly from 1996-1999. In some intervals as many as one third of all the news segments mention locations in Serbia.



## 3. ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under Cooperative Agreement No. IRI-9817496. Thanks to the Informedia group members for helpful discussions.

## 4. REFERENCES

- [1] Wactlar, H., Kanade, T., Smith, M., and Stevens, S. Intelligent Access to Digital Video: Informedia Project. *IEEE Computer* 29, 5 (1996), 46-52.
- [2] Derthick, M., Kolojechick, J.A., and Roth, S.F. [An interactive visual query environment for exploring data](#). In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '97)*, pages 189-198, 1997.